

DEC 13 2002

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Oon, Chong Jin
Lim, Gek Keow
Leong, Ai Lin
Zhao, Yi
Chen, Wei Ning
- (ii) TITLE OF INVENTION: A VACCINE-INDUCED HEPATITIS B VIRAL
STRAIN AND USES THEREOF
- (iii) NUMBER OF SEQUENCES: 11
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Ladas & Parry
 - (B) STREET: 26 West 61 Street
 - (C) CITY: New York
 - (D) STATE: New York
 - (E) COUNTRY: USA
 - (F) ZIP: 10023
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 09/719,533
 - (B) FILING DATE: 13-DEC-2000
 - (C) CLASSIFICATION: 435
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: PCT/SG98/00045
 - (B) FILING DATE: 19-JUN-1998
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Miller, Cynthia R.
 - (B) REGISTRATION NUMBER: 34,678
 - (C) REFERENCE/DOCKET NUMBER: U-013108-9
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (212) 708-1890

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3215 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCCACCACT TTCCACCAAA CTCTTCAAGA TCCCAGAGTC AGGGCCCTGT ACTTTCTCTG	
50	
TGGTGGCTCC AGTTCAGGAA CAGTGAGCCC TGCTCAGAAT ACTGTCTCTG CCATATCGTC	1
20	
AATCTTATCG AAGACTGGGG ACCCTGTACC GAACATGGAG AACATCGCAT CAGGACTCCT	1
80	
AGGACCCCTG CTCGTGTTAC AGGCGGGGTT TTTCTTGTTG ACAAAAATCC TCACAATACC	2
40	
GCAGAGTCTA GACTCGTGGT GGAATTCTCT CAATTTTCTA GGGGGAACAC CCGTGTGTCT	3
00	
TGGCCAAAAT TCGCAGTCCC AAATCTCCAG TCACTCACCA ACCTGTTGTC CTCCAATTTG	3
50	
TCCTGGTTAT CGCTGGATGT GTCTGCGGCG TTTTATCATE TTCTCTGCA TCCTGCTGCT	4
20	
ATGCCTCATC TTCTTGTTGG TTCTTCTGGA CTATCAAGGT ATGTTGCCCG TTTGTCTCT	4
80	
AATTCAGGA TCAACAACAA CCAGCACCGG ACCATGCAAA ACCTGCACAA CTCTGCTCA	5
40	
AGGAACCTCT ATGTTTCCCT CATGTTGCTG TACAAAACCT ACGGACAGAA ACTGCACCTG	6
00	
TATTCOCATC CCATCATCTT GGGCTTTCGC AAAATACCTA TGGGAGTGGG CCTCAGTCCG	6
60	
TTTCTCTTGG CTCAGTTTAC TAGTGCCATT TGTTCAGTGG TTCGTAGGGC TTTCCCCCAC	7
20	
TGTCTGGCTT TCAGTTATAT GGATGATGTG GTTTTGGGGG CCAAGTCTGT ACAACATCTT	7
80	
GAGTCCCTTT ATSCCGCTGT TACCAATTTT CTTTGTCTT TGGGTATACA TTTAAACCT	8
40	
CACAAAACAA AAAGATGGGG ATATTCCCTT AACTTCATGG GATATGTCAT TGGGAGTTGG	9
00	
GGCACATTGC CACAGGAACA TATTGTACAA AAAATCAAAA TGTGTTTTAG GAAACTTCCT	9
60	

GTAAACAGGC CTATTGATTG GAAAGTATGT CAACGAATTG TGGGTCTTTT GGGGTTTGCC 20	10
GCCCCCTTTCA CGCAATGTGG ATATCCTGCT TTAATGCCTT TATATGCATG TATACAAGCA 80	10
AAACAGGCTT TTACTTTCTC GCAAACCTAC AAGACCTTTC TAAGTAAACA GSTATCTGAAC 40	11
CTTTACCCCG TTGCTCGGCA ACGCCCTGGT CTGTGCCAAG TGTTTGCTGA CGCAACCCCG 00	12
ACTGGTTGGG GCTTGGGCAT AGGCCATCAG CGCATGCGTG GAACCTTTGT GTCTCCTCTG 60	12
CCGATCCATA CTGCGGAACT CCTAGCCGCT TGTTTTGCTC GCAGCAGGTC TGGGGCAAAA 20	13
CTCATCGGGA CTGACAATTC TGTCGTGCTC TCCCGCAAGT ATACATCATT TCCATGGCTG 80	13
CTAGGCTGTG CTGCCAACTG GATCCTGCGC GGGACGTCTT TTGTTTACGT CCGGTGCGCG 40	14
CTGAATCCCG CGSAGACCC CTCCCGGGGC CGCTTGGGGC TCTACCGCC GCTTCTCCGC 00	15
CTGTTATAAC GAGCGACAC GGGGCGCAC TCTCTTTACG CGGACTCCCC GTCTGTGCCT 60	15
TCTCATCTGC CGGACCTGTG GCACTTCGCT TCACCTCTGC ACGTCGCATG GAGACCACCG 20	16
TGAACGCCCA CGGGAACTG CCAAGGTCT TGCATAAGAG GACTCTTGGA CTTTCAGCAA 80	16
TGTCAACGAC CGACCTTGAG GCATACTTCA AAGACTGTGT GTTTAATGAG TGGGAGGAGT 40	17
TGGGGGAGGA GGTTAGGTTA AAGGTCTTTG TACTAGGAGG GTGTAGGCAT AAATTGGTGT 00	18
CTTCACCAFC ACCATGCAAC TTTTTCACCT CTGCCTAATC ATCTCATGTT CATGTCTAC 60	18
TGTTCAAGCC TCCAAGCTGT GCCTTGGGTG GCTTTGGGGC ATGGACATTG ACCCGTATAA 20	19
AGAATTTGGA GCTTCTGTGG AGTTACTCTC TTTTTGCTT TCTGACTTTT TTCTTCTAT 80	19
TCGAGATCTC CTCGACACCG CCTCTGCTCT GTATCGGGAG GCCTTAGAGT CTCCGGAACA 20	20

40

TTGTTCACT CACCATACGG CACTCAGGCA AGCTATTCTG AGTTGGGGTG AGTTAATGAA 00	21
TCTAGCCACC TGGGTGGGAA GTAATTTGGA AGATCCAGCA TCCAGGGAAT TAGTAGTCAG 60	21
CTATGTCAAC GTTAATATGG GCCTAAAAAT CAGACAACTA TTGTGGTTTC ACATTTCTCTG 20	22
TCTTACTTTT GGGAGAGAAA CTGTTCTTGA ATATTTGGTG TCTTTTGGAG TGTGGATTCTG 80	22
CACTCCTCCT GCATATAGAC CACCAAATGC CCCTATCTTA TCAACACTTC CGGAAACTAC 40	23
TGTTGTTAGA CGAAGAGGCA GGTCCCCTAG AAGAAGAACT CCCTCGCCTC GCAGACGAAG 00	24
GTCTCAATCG CCGCGTCGCA GAAGATCTCA ATCTCGGGAA TCTCAATGTT AGTATTCCTT 60	24
GGACACATAA GGTGGGAAAC TTTACGGGGC TTTATTCTTC TACGGTACCT TGCTTTAATC 20	25
CTAAATGGCA AACTCCTTCT TTTCCGGACA TTCATTGCA GGAGGACATT CTTGATAGAT 80	25
GTAAGCAATT TGTGGGGCCC CTTACAGTAA ATGAAAACAG GAGACTAAAA TTAATTATGC 40	26
CTGCTAGGTT TTATCCAAAT GTTACTAAAT ATTTGCCCTT AGATAAAGGG ATCAAACCAT 00	27
ATTATCCAGA GTATGTAGTT AATCATTACT TCCAGACGCG ACATTATTTA CACACTCTTT 60	27
GGAAGGCGGG GATCTTATAT AAAAGAGAGT CCACACGTAG CGCCTCATT TGGGGGTCAC 20	28
CATATTCTTG GGAACAAGAT CTACAGCATG GGAGGTTGGT CTTCCAAACC TCGAAAAGGC 80	28
ATGGGGACAA ATCTTTCTGT CCCCATTCCC CTGGGATTCT TCCCCGATCA TCAGTTGGAC 40	29
CCTGCATTCA AAGCCAACTC AGAAAATCCA GATTGGGACC TCAACCCGCA CAAGGACAAC 00	30
TGGCCGGACG CCAACAAGGT GGGAGTGGGA GCATTCGGGC CAGGGTTCAC CCCTCCTCAT 60	30

GGGGGACTGT TGGGGTGGAG CCCTCAGGCT CAGGGCCTAC TCACAACTGT GCCAGCAGCT 31
20

CCTCCTCCTG CCTCCACCAA TCGGCAGTCA GGAAGGCAGC CTACTCCCTT ATCTCCACCT 31
80

CTAAGGGACA CTCATCCTCA GGCCATGCAG TGGAA 32
15

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 843 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Pro	Leu	Ser	Tyr	Gln	His	Phe	Arg	Lys	Leu	Leu	Leu	Leu	Asp	Glu	
1				5					10					15		
Glu	Ala	Gly	Pro	Leu	Glu	Glu	Glu	Leu	Pro	Arg	Leu	Ala	Asp	Glu	Gly	
			20					25					30			
Leu	Asn	Arg	Arg	Val	Ala	Glu	Asp	Leu	Asn	Leu	Gly	Asn	Leu	Asn	Val	
		35					40					45				
Ser	Ile	Pro	Trp	Thr	His	Lys	Val	Gly	Asn	Phe	Thr	Gly	Leu	Tyr	Ser	
50						55					60					
Ser	Thr	Val	Pro	Cys	Phe	Asn	Pro	Lys	Trp	Gln	Thr	Pro	Ser	Phe	Pro	
65					70					75					80	
Asp	Ile	His	Leu	Gln	Glu	Asp	Ile	Leu	Asp	Arg	Cys	Lys	Gln	Phe	Val	
			85						90					95		
Glu	Pro	Leu	Thr	Val	Asn	Glu	Asn	Arg	Arg	Leu	Lys	Leu	Ile	Met	Pro	
			100					105					110			
Ala	Arg	Phe	Tyr	Pro	Asn	Val	Thr	Lys	Tyr	Leu	Pro	Leu	Asp	Lys	Gly	
		115					120					125				
Ile	Lys	Pro	Tyr	Tyr	Pro	Glu	Tyr	Val	Val	Asn	His	Tyr	Phe	Gln	Thr	
130						135					140					
Arg	His	Tyr	Leu	His	Thr	Leu	Trp	Lys	Ala	Gly	Ile	Leu	Tyr	Lys	Arg	
145					150					155					160	

Glu	Ser	Thr	Arg	Ser	Ala	Ser	Phe	Cys	Gly	Ser	Pro	Tyr	Ser	Trp	Glu
				165					170					175	
Gln	Asp	Leu	Gln	His	Gly	Arg	Leu	Val	Phe	Gln	Thr	Ser	Lys	Arg	His
			180					185					190		
Gly	Asp	Lys	Ser	Phe	Cys	Pro	Glu	Ser	Pro	Gly	Ile	Leu	Pro	Arg	Ser
		195					200					205			
Ser	Val	Gly	Pro	Cys	Ile	Gln	Ser	Gln	Leu	Arg	Lys	Ser	Arg	Leu	Gly
	210					215					220				
Pro	Gln	Pro	Ala	Gln	Gly	Gln	Leu	Ala	Gly	Arg	Gln	Gln	Gly	Gly	Ser
225					230					235					240
Gly	Ser	Ile	Arg	Ala	Arg	Val	His	Pro	Ser	Ser	Trp	Gly	Thr	Val	Gly
				245					250					255	
Val	Glu	Pro	Ser	Gly	Ser	Gly	Pro	Thr	His	Asn	Cys	Ala	Ser	Ser	Ser
			260					265					270		
Ser	Ser	Cys	Leu	His	Gln	Ser	Ala	Val	Arg	Lys	Ala	Ala	Tyr	Ser	Leu
		275					280					285			
Ile	Ser	Thr	Ser	Lys	Gly	His	Ser	Ser	Ser	Gly	His	Ala	Val	Glu	Leu
	290					295					300				
His	His	Phe	Pro	Pro	Asn	Ser	Ser	Arg	Ser	Gln	Ser	Gln	Gly	Pro	Val
305					310					315					320
Leu	Ser	Cys	Trp	Trp	Leu	Gln	Phe	Arg	Asn	Ser	Glu	Pro	Cys	Ser	Glu
				325					330					335	
Tyr	Cys	Leu	Cys	His	Ile	Val	Asn	Leu	Ile	Glu	Asp	Trp	Gly	Pro	Cys
			340					345					350		
Thr	Glu	His	Gly	Glu	His	Arg	Ile	Arg	Thr	Pro	Arg	Thr	Pro	Ala	Arg
		355					360					365			
Val	Thr	Gly	Gly	Val	Phe	Leu	Val	Asp	Lys	Asn	Pro	His	Asn	Thr	Ala
	370					375					380				
Glu	Ser	Arg	Leu	Val	Val	Asp	Phe	Ser	Gln	Phe	Ser	Arg	Gly	Asn	Thr
385					390					395					400
Arg	Val	Ser	Trp	Pro	Lys	Phe	Ala	Val	Pro	Asn	Leu	Gln	Ser	Leu	Thr
				405					410					415	
Asn	Leu	Leu	Ser	Ser	Asn	Leu	Ser	Trp	Leu	Ser	Leu	Asp	Val	Ser	Ala
			420					425					430		

Ala	Phe	Tyr	His	Leu	Pro	Leu	His	Pro	Ala	Ala	Met	Pro	His	Leu	Leu	435	440	445
Val	Gly	Ser	Ser	Gly	Leu	Ser	Arg	Tyr	Val	Ala	Arg	Leu	Ser	Ser	Asn	450	455	460
Ser	Arg	Ile	Asn	Asn	Asn	Glu	His	Arg	Thr	Met	Glu	Asn	Leu	His	Asn	465	470	475
Ser	Cys	Ser	Arg	Asn	Leu	Tyr	Val	Ser	Leu	Met	Leu	Leu	Tyr	Lys	Thr	485	490	495
Tyr	Gly	Gln	Lys	Leu	His	Leu	Tyr	Ser	His	Pro	Ile	Ile	Leu	Gly	Phe	500	505	510
Arg	Lys	Ile	Pro	Met	Gly	Val	Gly	Leu	Ser	Pro	Phe	Leu	Leu	Ala	Gln	515	520	525
Phe	Thr	Ser	Ala	Ile	Cys	Ser	Val	Val	Arg	Arg	Ala	Phe	Pro	His	Cys	530	535	540
Leu	Ala	Phe	Ser	Tyr	Met	Asp	Asp	Val	Val	Leu	Gly	Ala	Lys	Ser	Val	545	550	555
Gln	His	Leu	Glu	Ser	Leu	Tyr	Ala	Ala	Val	Thr	Asn	Phe	Leu	Leu	Ser	565	570	575
Leu	Gly	Ile	His	Leu	Asn	Pro	His	Lys	Thr	Lys	Arg	Trp	Gly	Tyr	Ser	580	585	590
Leu	Asn	Phe	Met	Gly	Tyr	Val	Ile	Gly	Ser	Trp	Gly	Thr	Leu	Pro	Gln	595	600	605
Glu	His	Ile	Val	Gln	Lys	Ile	Lys	Met	Cys	Phe	Arg	Lys	Leu	Pro	Val	610	615	620
Asn	Arg	Pro	Ile	Asp	Trp	Lys	Val	Cys	Gln	Arg	Ile	Val	Gly	Leu	Leu	625	630	635
Gly	Phe	Ala	Ala	Pro	Phe	Thr	Gln	Cys	Gly	Tyr	Pro	Ala	Leu	Met	Pro	645	650	655
Leu	Tyr	Ala	Cys	Ile	Gln	Ala	Lys	Gln	Ala	Phe	Thr	Phe	Ser	Gln	Thr	660	665	670
Tyr	Lys	Thr	Phe	Leu	Ser	Lys	Gln	Tyr	Leu	Asn	Leu	Tyr	Pro	Val	Ala	675	680	685
Arg	Gln	Arg	Pro	Gly	Leu	Cys	Glu	Val	Phe	Ala	Asp	Ala	Thr	Pro	Thr	690	695	700

Gly	Trp	Gly	Leu	Ala	Ile	Gly	His	Gln	Arg	Met	Arg	Gly	Thr	Phe	Val	705	710	715	720
Ser	Pro	Leu	Pro	Ile	His	Thr	Ala	Glu	Leu	Leu	Ala	Ala	Cys	Phe	Ala	725	730		735
Arg	Ser	Arg	Ser	Gly	Ala	Lys	Leu	Ile	Gly	Thr	Asp	Asn	Ser	Val	Val	740	745		750
Leu	Ser	Arg	Lys	Tyr	Thr	Ser	Phe	Pro	Trp	Leu	Leu	Gly	Cys	Ala	Ala	755	760		765
Asn	Trp	Ile	Leu	Arg	Gly	Thr	Ser	Phe	Val	Tyr	Val	Pro	Ser	Ala	Leu	770	775		780
Asn	Pro	Ala	Asp	Asp	Pro	Ser	Arg	Gly	Arg	Leu	Gly	Leu	Tyr	Arg	Pro	785	790		795
Leu	Leu	Arg	Leu	Leu	Tyr	Arg	Pro	Thr	Thr	Gly	Arg	Thr	Ser	Leu	Tyr	805	810		815
Ala	Asp	Ser	Pro	Ser	Val	Pro	Ser	His	Leu	Pro	Asp	Arg	Val	His	Phe	820	825		830
Ala	Ser	Pro	Leu	His	Val	Ala	Trp	Arg	Pro	Pro						835	840		

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 400 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Gly	Gly	Trp	Ser	Ser	Lys	Pro	Arg	Lys	Gly	Met	Gly	Thr	Asn	Leu	1	5	10	15
Ser	Val	Pro	Asn	Pro	Leu	Gly	Phe	Phe	Pro	Asp	His	Gln	Leu	Asp	Pro	20	25		30
Ala	Phe	Lys	Ala	Asn	Ser	Glu	Asn	Pro	Asp	Trp	Asp	Leu	Asn	Pro	His	35	40		45
Lys	Asp	Asn	Trp	Pro	Asp	Ala	Asn	Lys	Val	Gly	Val	Gly	Ala	Phe	Gly				

50					55					60					
Pro 65	Gly	Phe	Thr	Pro	Pro 70	His	Gly	Gly	Leu	Leu 75	Gly	Trp	Ser	Pro	Gln 80
Ala	Gln	Gly	Leu	Leu 85	Thr	Thr	Val	Pro	Ala 90	Ala	Pro	Pro	Pro	Ala 95	Ser
Thr	Asn	Arg	Gln 100	Ser	Gly	Arg	Gln	Pro 105	Thr	Pro	Leu	Ser	Pro 110	Pro	Leu
Arg	Asp	Thr 115	His	Pro	Gln	Ala	Met 120	Gln	Trp	Asn	Ser	Thr	Thr 125	Phe	His
Gln 130	Thr	Leu	Gln	Asp	Pro	Arg 135	Val	Arg	Ala	Leu	Tyr 140	Phe	Pro	Ala	Gly
Gly 145	Ser	Ser	Ser	Gly	Thr 150	Val	Ser	Pro	Ala	Gln 155	Asn	Thr	Val	Ser	Ala 160
Ile	Ser	Ser	Ile	Leu 165	Ser	Lys	Thr	Gly	Asp 170	Pro	Val	Pro	Asn	Met 175	Glu
Asn	Ile	Ala	Ser 180	Gly	Leu	Leu	Gly	Pro 185	Leu	Leu	Val	Leu	Gln 190	Ala	Gly
Phe	Phe	Leu 195	Leu	Thr	Lys	Ile	Leu 200	Thr	Ile	Pro	Gln	Ser 205	Leu	Asp	Ser
Trp	Trp	Thr 210	Ser	Leu	Asn	Phe 215	Leu	Gly	Gly	Pro	Thr 220	Val	Cys	Leu	Gly
Gln 225	Asn	Ser	Gln	Ser	Gln 230	Ile	Ser	Ser	His	Ser 235	Pro	Thr	Cys	Cys	Pro 240
Pro	Ile	Cys	Pro	Gly 245	Tyr	Arg	Trp	Met	Cys 250	Leu	Arg	Arg	Phe	Ile 255	Ile
Phe	Leu	Cys	Ile 260	Leu	Leu	Leu	Cys	Leu 265	Ile	Phe	Leu	Leu	Val 270	Leu	Leu
Asp	Tyr	Gln 275	Gly	Met	Leu	Pro	Val 280	Cys	Pro	Leu	Ile	Pro 285	Gly	Ser	Thr
Thr 290	Thr	Ser	Thr	Gly	Pro	Cys 295	Lys	Thr	Cys	Thr	Thr 300	Pro	Ala	Gln	Gly
Thr 305	Ser	Met	Phe	Pro	Ser 310	Cys	Cys	Cys	Thr	Lys 315	Pro	Thr	Asp	Arg	Asn 320
Cys	Thr	Cys	Ile	Pro 325	Ile	Pro	Ser	Ser	Trp 330	Ala	Phe	Ala	Lys	Tyr 335	Leu

Trp	Glu	Trp	Ala	Ser	Val	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	Pro
			340					345					350		
Phe	Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	Val
		355					360					365			
Ile	Trp	Met	Met	Trp	Phe	Trp	Gly	Pro	Ser	Leu	Tyr	Asn	Ile	Leu	Ser
	370					375					380				
Pro	Phe	Met	Pro	Leu	Leu	Pro	Ile	Phe	Phe	Cys	Leu	Trp	Val	Tyr	Ile
385					390					395					400

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 212 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Gln	Leu	Phe	His	Leu	Cys	Leu	Ile	Ile	Ser	Cys	Ser	Cys	Pro	Thr
1				5					10					15	
Val	Gln	Ala	Ser	Lys	Leu	Cys	Leu	Gly	Trp	Leu	Trp	Gly	Met	Asp	Ile
			20					25					30		
Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Ser	Val	Glu	Leu	Leu	Ser	Phe	Leu
		35					40					45			
Pro	Ser	Asp	Phe	Phe	Pro	Ser	Ile	Arg	Asp	Leu	Leu	Asp	Thr	Ala	Ser
	50					55					60				
Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	Glu	His	Cys	Ser	Pro	His
65					70					75					80
His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Ser	Trp	Gly	Glu	Leu	Met	Asn
				85					90					95	
Leu	Ala	Thr	Trp	Val	Gly	Ser	Asn	Leu	Glu	Asp	Pro	Ala	Ser	Arg	Glu
			100					105					110		
Leu	Val	Val	Ser	Tyr	Val	Asn	Val	Asn	Met	Gly	Leu	Lys	Ile	Arg	Gln
		115					120					125			
Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	Phe	Gly	Arg	Glu	Thr	Val
	130					135					140				

Leu	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	Pro	Pro	Ala
145					150					155					160
Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	Glu	Thr	Thr
				165					170					175	
Val	Val	Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro
			180					185					190		
Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Arg
		195					200					205			
Glu	Ser	Gln	Cys												
	210														

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 154 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ala	Ala	Arg	Leu	Cys	Cys	Gln	Leu	Asp	Pro	Ala	Arg	Asp	Val	Leu
1				5					10					15	
Cys	Leu	Arg	Pro	Val	Gly	Ala	Glu	Ser	Arg	Gly	Arg	Pro	Leu	Pro	Gly
			20					25					30		
Pro	Leu	Gly	Ala	Leu	Pro	Pro	Ala	Ser	Pro	Pro	Val	Ile	Pro	Thr	Asp
		35					40					45			
His	Gly	Ala	His	Leu	Ser	Leu	Arg	Gly	Leu	Pro	Val	Cys	Ala	Phe	Ser
	50					55					60				
Ser	Ala	Gly	Pro	Cys	Ala	Leu	Arg	Phe	Thr	Ser	Ala	Arg	Arg	Met	Glu
65				70						75				80	
Thr	Thr	Val	Asn	Ala	His	Gly	Asn	Leu	Pro	Lys	Val	Leu	His	Lys	Arg
			85					90					95		
Thr	Leu	Gly	Leu	Ser	Ala	Met	Ser	Thr	Thr	Asp	Leu	Glu	Ala	Tyr	Phe
		100					105					110			
Lys	Asp	Cys	Val	Phe	Asn	Glu	Trp	Glu	Glu	Leu	Gly	Glu	Glu	Val	Arg
	115					120						125			
Leu	Lys	Val	Phe	Val	Leu	Gly	Gly	Cys	Arg	His	Lys	Leu	Val	Cys	Ser

130

135

140

Pro Ser Pro Cys Asn Phe Phe Thr Ser Ala
145 150

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATAAGCTTAT GCCCCTATCT TATCAACACT TCCGGA
36

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAGTCTAGAC TCTGCGGTAT TGTGA
25

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAGTCTAGAC TCGTGGTGGG CTTCT
25

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGAGAATTCT CACGGTGGTC TCCATGCGAC GT
32

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTTGTTTACG TCCCGT
16

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTAAGCTTAG TTTCCGGAAG TGTTGAT
27